



PIBIC/CNPq/UFPG-2010

## **ANÁLISE E CARACTERIZAÇÃO DE SINAIS DE VOZ COM PATOLOGIAS NA LARINGE**

**Camila Duarte Rodrigues Lopes<sup>1</sup>, Benedito Guimarães Aguiar Neto<sup>2</sup>**

### **RESUMO**

O principal objetivo deste trabalho foi a avaliação de vozes patológicas através das técnicas de processamento digital de sinais, tendo como foco a extração das características acústicas de longa duração e sua análise no domínio do tempo. Foram utilizados arquivos de vozes de um banco de dados conhecido - Disordered Voice Database Model 4337. A extração das características foi realizada em um computador pessoal pelos softwares *Multi-Speech – Signal Analysis Workstation*, Modelo 3700 e *Multi-Dimensional Voice Program (MDVP)*, Modelo 5105, da Kay Elemetrics. Tanto vozes normais (pessoas saudáveis), como vozes patológicas foram utilizadas. Os valores obtidos foram organizados em três grupos: vozes normais, vozes patológicas com edema e vozes patológicas com pólipos. Além da diferenciação do sexo do paciente. Os resultados sugerem o uso em conjunto das várias características, para a diferenciação entre voz normal e patológica.

**Palavras-chave:** Análise de voz, Análise acústica, estimação de pitch

### **ANALYSIS AND CHARACTERIZATION OF VOICE SIGNALS WITH PATHOLOGIES IN LARYNX**

#### **ABSTRACT**

The main objective was the evaluation of pathological voices using the techniques of digital signal processing, focusing on the extraction of the long duration acoustic characteristics and its analysis in time domain. Files used were from a well known database - Disordered Voice Database Model 4337. The extraction of the characteristics was performed on a personal computer by the softwares *Multi-Speech - Signal Analysis Workstation* Model 3700 and *Multi-Dimensional Voice Program (MDVP)*, Model 5105, Kay Elemetrics. Both normal (healthy) and pathological voices were used. The values were organized into three groups: normal voices, pathological voices with edema and pathological voices with polyps. And also the sex differentiation of the patient was taken into account. The results suggest the combined use of several characteristics, for differentiation between normal and pathological voice.

**Keywords:** Voice analysis, Acoustic analysis, pitch estimation

### **INTRODUÇÃO**

Técnicas de processamento digital de sinais de voz, aplicadas à detecção de patologias, que afetam de alguma forma o aparelho fonador, se mostram em constante desenvolvimento devido à praticidade

<sup>1</sup> Aluna do Curso de Engenharia Elétrica, Unidade Acadêmica de Engenharia Elétrica, UFPG, Campina Grande, PB, E-mail: camila.lobes@ee.ufcg.edu.br

<sup>2</sup> Engenheiro Eletricista, Professor. Doutor, Unidade Acadêmica de Engenharia Elétrica, UFPG, Campina Grande, PB, E-mail: bganeto@dee.ufcg.edu.br; \*Autor para correspondências.

apresentada por essas técnicas. Diferentemente de exames realizados através da observação direta dos órgãos responsáveis pela produção da fala - os quais necessitam do uso de aparelhos de custo elevado - além de serem invasivos e muitas vezes incômodos, a avaliação com o processamento da voz do paciente é uma ferramenta atrativa por seu baixo custo e facilidade no manuseio e obtenção das informações necessárias.

Patologias da laringe como nódulos nas cordas vocais, pólipos, cistos, carcinomas e paralisia dos nervos laríngeos, por exemplo, podem ser corrigidos por meio de: terapia vocal, cirurgia e, em alguns casos, radioterapia (MARTINEZ & RUFINER, 2000). Entretanto, antes dessas intervenções se faz necessário avaliar o nível e o tipo da patologia. Outro aspecto importante da avaliação da voz por meio do processamento digital é a possibilidade do acompanhamento da evolução do paciente, submetido a tratamentos como os citados, sem a necessidade da repetição de exames invasivos.

Uma das abordagens utilizadas para a avaliação da voz do paciente é aquela que leva em consideração as medidas clássicas para análise acústica. Dentre essas medidas, o *pitch* se destaca como uma das principais características.

A frequência fundamental ( $f_0$ ) de um sinal, se refere à componente de maior amplitude e de menor frequência dentre suas harmônicas. Para sinais acústicos, as definições de frequência fundamental e *pitch* se confundem e muitos autores tratam essas duas características acústicas como uma só. Rigorosamente, o *pitch* de um sinal acústico é uma medida perceptual, tomada a partir da resposta do aparelho auditivo humano quando excitado por um determinado som. Em um sinal de voz, a frequência fundamental corresponde à frequência do sinal de excitação proveniente da vibração das cordas vocais. Como as relações entre as escalas de  $f_0$  e do *pitch* são muito bem conhecidas, essa distinção se torna irrelevante (GERHARD, 2003) e ambos os termos são usados com um mesmo significado. Assim, é comum referir-se a *período de pitch* como uma medida do período da frequência fundamental.

Para o projeto de pesquisa proposto, o *pitch* do sinal de voz foi escolhido como a primeira característica acústica a ser analisada pela sua importância e também pela dependência que outras medidas acústicas clássicas, como *jitter* e *shimmer*, apresentam em relação ao mesmo.

As patologias na laringe afetam as cordas vocais de maneira a modificar sua resposta à passagem do ar proveniente dos pulmões, afetando a frequência de sua vibração. É justamente essa mudança, identificada a partir da variação do *pitch*, que constitui a base da análise e caracterização de sinais de voz patológica, a partir dos parâmetros de longa duração.

Na segunda etapa do projeto, alguns parâmetros utilizados para caracterizar acusticamente o sinal de voz a partir da frequência fundamental foram analisados. Foram eles: *jitter*, *shimmer*, quociente de perturbação de amplitude e quociente de perturbação do *pitch*. Além destes, outros parâmetros foram igualmente utilizados: a relação harmônica-ruído (RHR) e o índice de turbulência vocal (ITV). Estas características de um sinal de voz formam os chamados parâmetros acústicos a longos intervalos de tempo, obtidos a partir da média das perturbações locais medidas (COSTA, 2008).

O *Jitter* é uma medida de avaliação da variabilidade período-a-período do período de *pitch* dentro da amostra de voz analisada. *Shimmer* é a avaliação em variabilidade da amplitude pico-a-pico dentro da amostra de voz analisada.

O quociente de perturbação de amplitude ou APQ (*Amplitude Perturbation Quotient*) é obtido através da avaliação relativa da variabilidade período-período da amplitude de pico-a-pico dentro da amostra de voz analisada, ou seja, mede a irregularidade da amplitude de pico-a-pico do sinal de voz.

O quociente de Perturbação Periódica do *Pitch*, PPQ (*Pitch Period Perturbation Quotient*) representa a avaliação relativa da variabilidade do *pitch* dentro da amostra de voz analisada.

A relação harmônica-ruído (RHR) é a relação média entre a energia espectral intra-harmônicas (*inharmonics*) na faixa de frequência de 1500-4500 Hz e a energia espectral das harmônicas na faixa de frequência de 70-4500 Hz. É uma avaliação geral do ruído presente no sinal de voz analisado.

O Índice de Turbulência Vocal, VTI (*Voice Turbulence Index*) mede a taxa média da energia espectral das altas frequências, na faixa de 2800-5800 Hz em áreas do sinal onde a influência da frequência e variações de amplitude, paradas vocais e componentes sub-harmônicas é mínima. Em resumo, o VTI mede o nível de energia relativa do ruído de alta frequência.

Com os valores obtidos, uma análise estatística e comparativa entre os valores normais e patológicos permite a classificação das vozes.

## METODOLOGIA

Para esta pesquisa, as principais ferramentas de trabalho utilizadas foram os softwares *Multi-Speech – Signal Analysis Workstation*, Modelo 3700 e o Multi-Dimensional Voice Program (MDVP), Model 5105, da Kay Elemetrics, USA (KAY ELEMETRICS, 1994)

A extração do *pitch* foi feita no *Multi-Speech*, enquanto o restante dos valores foram obtidos utilizando a análise do MDVP.

Os dados utilizados foram gravados pelo *Massachusetts Eye and Ear Infirmary (MMEI) Voice and Speech Lab (KAY ELEMETRICS, 1994)*. Este banco de dados de vozes desordenadas (*Disordered Voice database – Model 4337*) inclui mais de 1.400 arquivos de voz gravados em CD-ROM (vogal /a/ sustentada e 12 segundos da *Rainbow Passage*) de aproximadamente 700 locutores. Os dados foram gravados em um ambiente controlado e amostrados a taxas de 25 ou 50 k amostras/s, com resolução de 16 bits/amostra. Devido ao interesse deste trabalho ser focado em patologias nas cordas vocais, foram utilizados somente arquivos da vogal /a/ sustentada. Dado que a produção dos sons de vogais sustentadas decorre da vibração das cordas vocais, qualquer alteração neste padrão pode ser melhor evidenciado para tais caso. Ao todo, foram 53 arquivos de vozes normais – 32 femininas e 21 masculinas, 44 arquivos de vozes afetadas por Edema – 33 femininas e 11 masculinas e 20 de vozes afetadas por Pólipos – 8 femininas e 12 masculinas.

Tabela 1: Arquivos de vozes normais

| Nome do arquivo |         |                  |         |
|-----------------|---------|------------------|---------|
| Vozes Femininas |         | Vozes Masculinas |         |
| AXHNAL          | LAD1NAL | BJB1NAL          | RJS1NAL |
| BJV1NAL         | LDP1NAL | DJG1NAL          | SIS1NAL |
| CAD1NAL         | LLA1NAL | DWS1NAL          | SXV1NAL |
| CEB1NAL         | LMV1NAL | EJC1NAL          | TXN1NAL |
| DAJ1NAL         | LMW1NAL | FMB1NAL          | WDK1NAL |
| DFP1NAL         | MAM1NAL | GPC1NAL          |         |
| DMA1NAL         | MCB1NAL | GZZ1NAL          |         |
| EDC1NAL         | MXB1NAL | JMC1NAL          |         |
| HBL1NAL         | MXZ1NAL | KAN1NAL          |         |
| JAF1NAL         | NJS1NAL | MAS1NAL          |         |
| JAN1NAL         | PBD1NAL | MFM1NAL          |         |
| JAP1NAL         | SCK1NAL | MJU1NAL          |         |
| JEG1NAL         | SCT1NAL | OVK1NAL          |         |
| JKR1NAL         | SEB1NAL | PCA1NAL          |         |
| JTH1NAL         | SLC1NAL | RHG1NAL          |         |
| JXC1NAL         | VMC1NAL | RHM1NAL          |         |

Tabela 2: Arquivos de vozes patológicas afetadas por EDEMA

| Nome do arquivo |         |                  |
|-----------------|---------|------------------|
| Vozes Femininas |         | Vozes Masculinas |
| ANA15AN         | KLC09AN | CTB30AN          |
| ANB28AN         | LAC02AN | DMG07AN          |
| CAC10AN         | LAD13AN | DXC22AN          |
| CAK25AN         | LGM01AN | JJD29AN          |
| CER16AN         | LXC01AN | JXB16AN          |
| DBF18AN         | LXD22AN | OAB28AN          |
| DJF23AN         | MCA07AN | PAT10AN          |
| EED07AN         | MCW21AN | RJL28AN          |
| EXE06AN         | NFG08AN | RTL27AN          |
| HLM24AN         | NLC08AN | RXP02AN          |
| JAJ31AN         | PMF03AN | WST20AN          |
| JMC18AN         | RCC11AN |                  |
| JMH22AN         | SLC23AN |                  |
| JXC21AN         | SXG23AN |                  |
| JXF11AN         | TLP13AN |                  |
| JXS09AN         | VAW07AN |                  |
| KAB03AN         |         |                  |

Tabela 3: Arquivos de vozes patológicas afetadas por PÓLIPOS

| Nome do arquivo |         |                  |         |
|-----------------|---------|------------------|---------|
| Vozes Femininas |         | Vozes Masculinas |         |
| CXR13AN         | MPB23AN | DAP17AN          | MPF25AN |
| JMH22AN         | SAV18AN | DMG24AN          | RHP12AN |
| KCG25AN         |         | EXI04AN          | RJR15AN |
| MCB20AN         |         | HWR04AN          | RWC23AN |
| MLF13AN         |         | JBS17AN          | SJD28AN |
| MMS29AN         |         | JLS11AN          | WPB30AN |

Os arquivos foram divididos em segmentos de 20 ms, com o avanço de 10 ms da janela de análise. Com isso, a sobreposição é de 50%.

Por meio dos *softwares* citados, uma série de valores numéricos é gerada, juntamente com o gráfico do contorno de *pitch* ao longo dos intervalos de análise. No caso da extração de *pitch*, foi escolhido o valor denominado *mean f<sub>o</sub>* (*f<sub>o</sub> média*), apenas pelo fácil entendimento de sua denominação.

A análise do MDVP gera uma série de gráficos, valores numéricos e um diagrama ou gráfico radial, onde são expostos os valores obtidos para o arquivo em análise em comparação com valores limites estabelecidos.

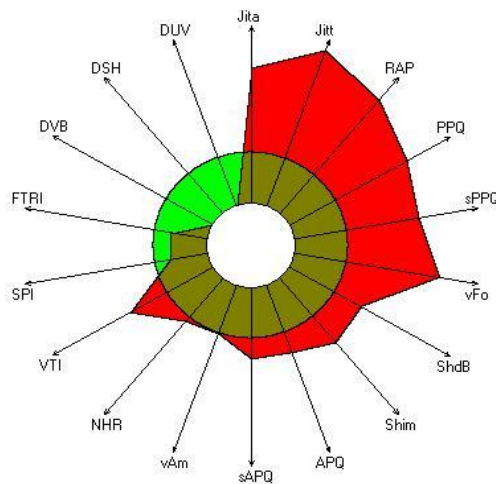


Figura 1. Exemplo de gráfico radial para arquivo de voz patológica, com edema unilateral, vogal sustentada /a/.

Apesar de o programa gerar cerca de 20 valores para diferentes características, foram utilizadas apenas as seguintes: o Jitter absoluto ou Jita, o Jitter percentual ou Jitt, Shimmer (dB) ShdB, Shimmer (Shim) em percentual, quociente de perturbação de amplitude (%) ou APQ, quociente de perturbação periódica do *pitch* PPQ, relação harmônica-ruído NHR e o índice de turbulência vocal VTI

Após a obtenção dos valores, eles foram organizados em tabelas para uma melhor visualização. Valores estatísticos como o valor médio e o desvio padrão dentro de cada grupo analisado foram calculados para que pudessem ser melhor expostos os resultados e conclusões.

Utilizando o *Microsoft Office Excel 2007* foi feita a exposição dos resultados em forma de gráficos comparativos. No caso do *pitch* os gráficos foram plotados em dois tipos: Voz Normal *versus* Voz Patológica (Edema ou Pólipo) e também de maneira a comparar os arquivos com valores máximos e mínimos em cada grupo de análise.

Para a análise utilizando os outros parâmetros, foram plotados gráficos para cada parâmetro no caso de arquivos de vozes femininas e masculinas.

Uma observação deve ser feita: alguns arquivos de vozes afetadas por patologias apresentam um grau de seriedade da doença bastante severo, o que resulta em uma voz altamente ruidosa. Tal característica dificulta, e às vezes torna impossível, a tarefa de extração dos parâmetros acústicos. Por esse motivo, não foi possível realizar a análise em alguns arquivos de voz patológica.

## RESULTADOS E DISCUSSÃO

Os resultados da primeira etapa da pesquisa realizada dizem respeito à frequência fundamental  $f_0$  e à sua influência da distinção entre vozes normais e patológicas, bem como numa tentativa de diferenciação entre as patologias do aparelho fonador analisadas – Edema e Pólipos nas dobras vocais.

Com o valor da  $f_0$  média para cada um dos arquivos já citados, os gráficos foram organizados da seguinte forma, de acordo com a natureza do mesmo:

- NORMAL x EDEMA – Feminino e Masculino;
- NORMAL x PÓLIPO – Feminino e Masculino,
- EDEMA x PÓLIPO – Feminino e Masculino.

Algumas observações devem ser feitas: em alguns casos, se o número de amostras é bem diferente, deve-se então dar uma maior atenção aos valores do que apenas à aparência dos gráficos; vale salientar que os valores foram plotados de maneira aleatória de acordo com o número de arquivos e a ordem alfabética dos seus nomes.

Para o grupo feminino, observou-se que na média geral, a diferenciação entre voz patológica e normal acontece pelos valores de  $f_0$ . Porém, em muitos casos, alguns arquivos de voz podem ser erroneamente classificados se somente for levado em consideração seu valor de frequência fundamental.

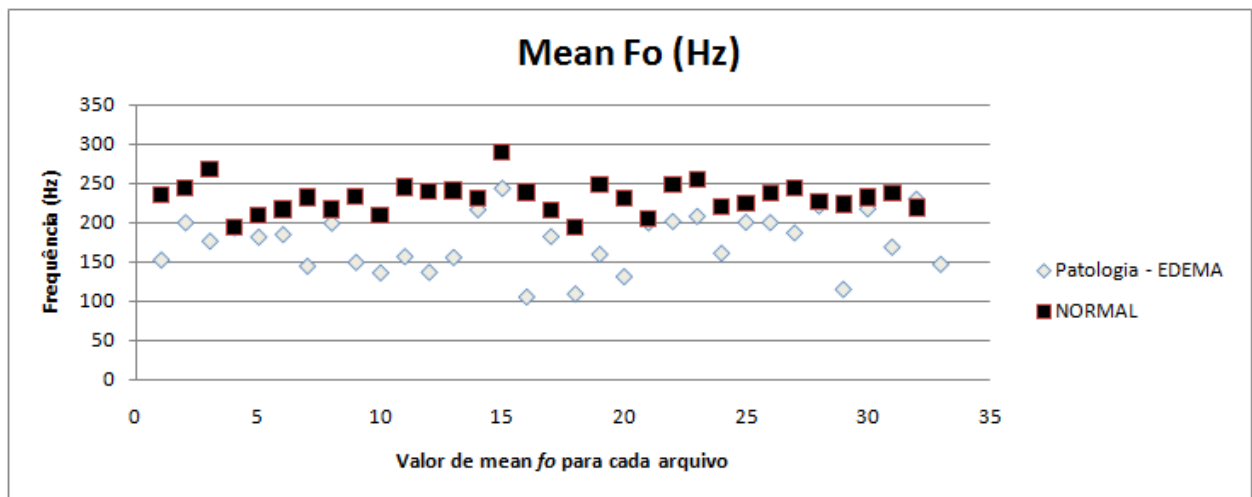


Figura 2. Valores médios para a frequência fundamental - Normal e EDEMA (Vozes Femininas).

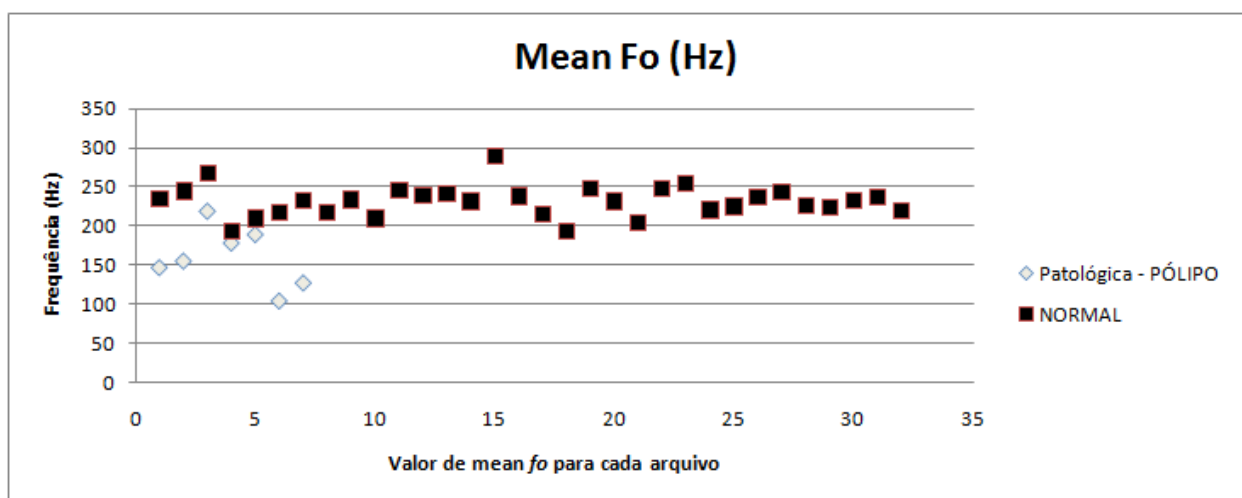


Figura 3. Valores médios para a frequência fundamental - Normal e PÓLIPO (Vozes Femininas).

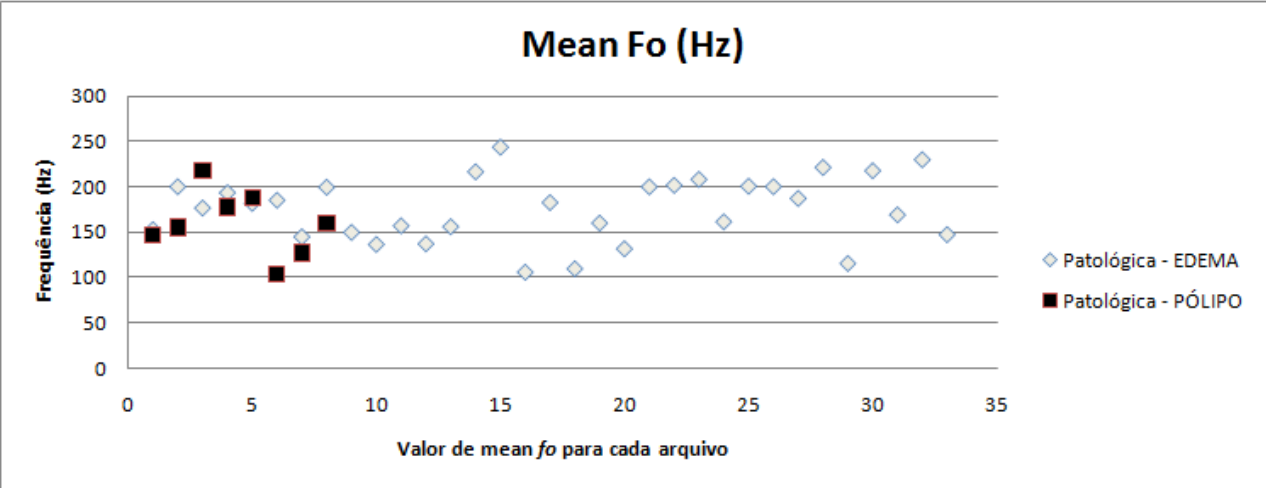


Figura 4. Valores médios para a frequência fundamental - EDEMA e PÓLIPOS (Vozes Femininas).

No caso do grupo masculino, um comportamento semelhante é evidenciado. Para diferenciar entre normal e patológica, a frequência fundamental pode auxiliar sim, mas não se mostrou suficiente. Já no caso de distinção entre as patologias, não foi observada diferenças nos valores obtidos, ou seja, não foi possível classificar o tipo de patologia pela frequência fundamental extraída.

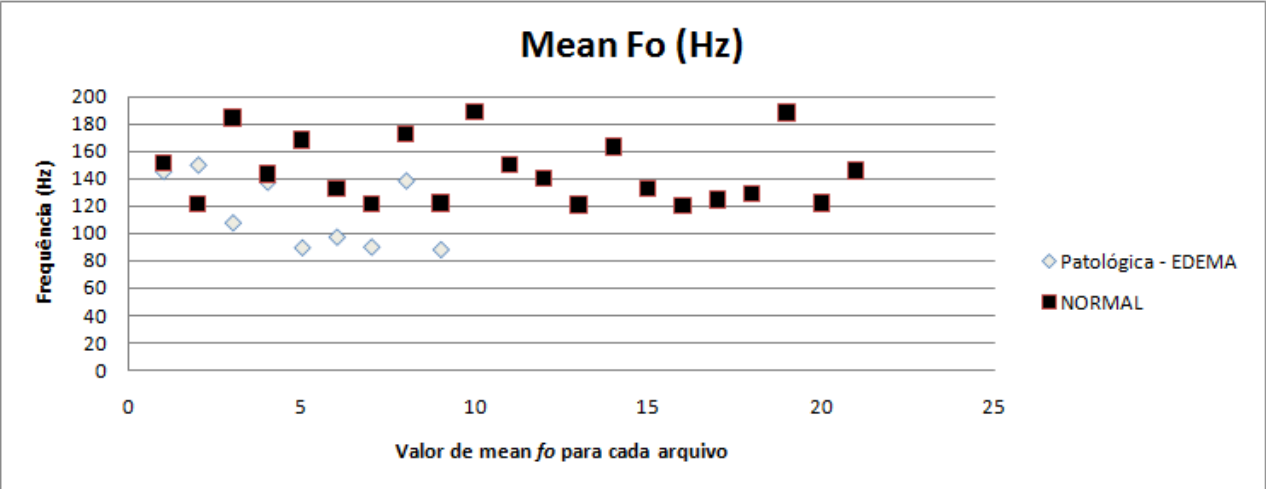


Figura 5. Valores médios para a frequência fundamental - Normal e EDEMA (Vozes Masculinas).

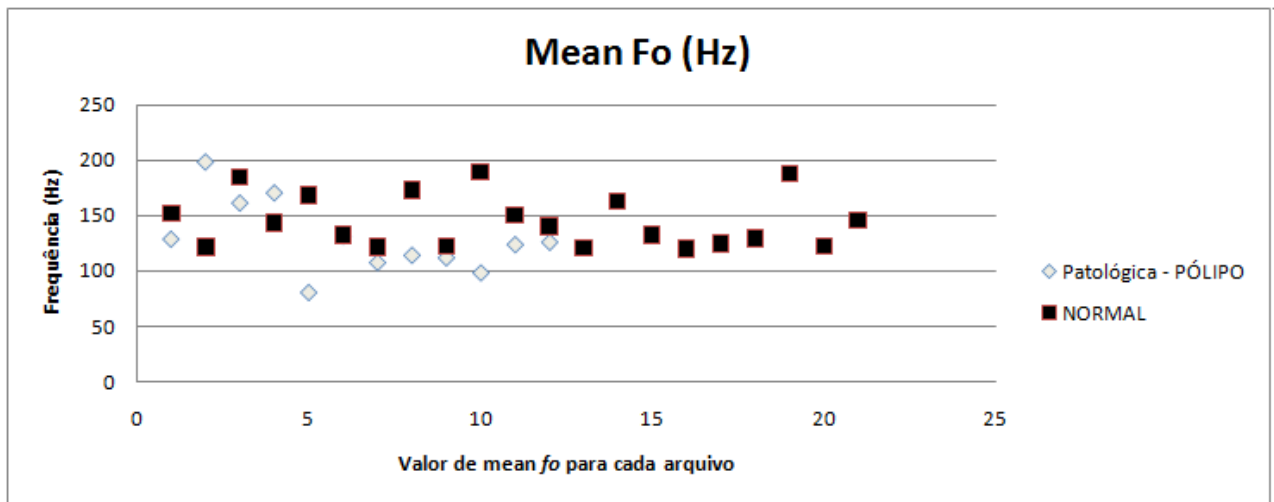


Figura 6. Valores médios para a frequência fundamental - Normal e PÓLIPOS (Vozes Masculinas).

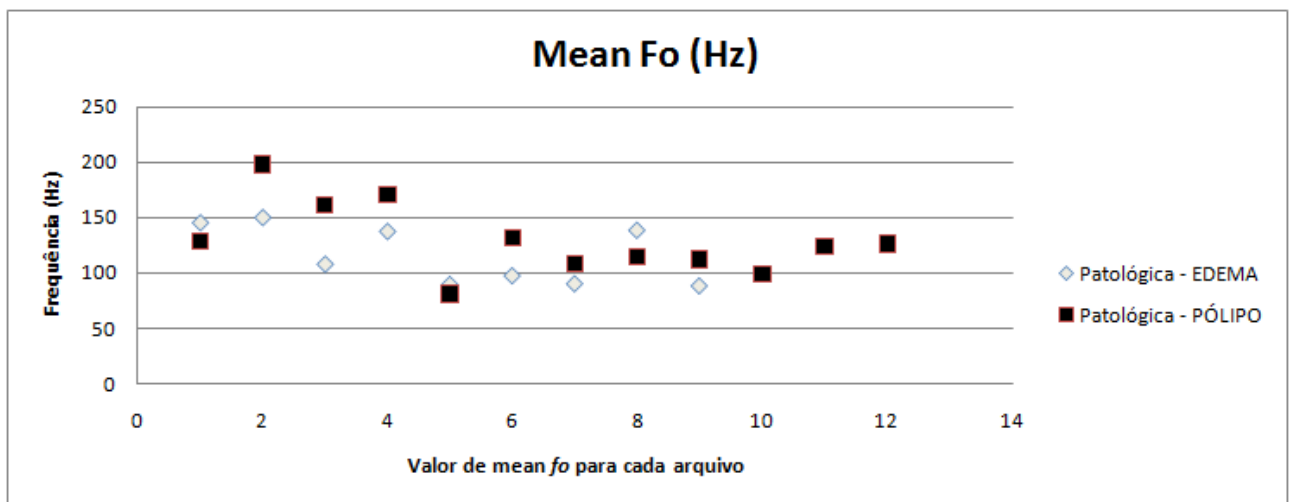


Figura 7. Valores médios para a frequência fundamental - EDEMA e PÓLIPOS (Vozes Masculinas).

É perceptível que a média geral do valor de *pitch* de uma voz patológica é menor que o de vozes de pessoas saudáveis. Tanto para o caso de pacientes masculinos quanto para femininos, foi mostrada essa diferença através da observação da faixa de variação de valores, bem como do valor médio tomando os dados de todos os arquivos (ver tabela 4). Além disso, outro valor analisado é o Coeficiente de Variação ( $C_v$ ), definido como uma medida de dispersão que possibilita a comparação de distribuições diferentes. Ou seja, mostra o quanto o desvio padrão tem peso sobre os valores apresentados; quanto menor o  $C_v$  mais homogêneo é o conjunto de dados. Seguindo este raciocínio, percebe-se também os maiores valores nos casos de vozes patológicas.

Tabela 5: Valor médio para o  $f_0$  nos conjuntos de arquivos

| Grupo                   | Média dos valores $f_0$ (Hz) | Desvio Padrão $\sigma$ | Coefficiente de Variação<br>( $c_v = \frac{\sigma}{m\u00e9dia}$ ) x 100 |
|-------------------------|------------------------------|------------------------|---|
| Normal – Feminina       | 231,25                       | 19,85                  | 8,58%   |
| Normal – Masculina      | 144,96                       | 23,91                  | 16,49%  |
| Edema – Feminina        | 174,55                       | 35,55                  | 20,37%  |
| Edema – Masculina       | 115,82                       | 26,07                  | 22,51%  |
| P\u00f3lipo – Feminina  | 159,66                       | 38,51                  | 24,12%  |
| P\u00f3lipo - Masculina | 129,48                       | 32,55                  | 25,14%  |

Para evidenciar que um diagn\u00f3stico de voz patol\u00f3gica baseado somente na frequ\u00eancia fundamental deve ser evitado, foi realizada a an\u00e1lise mais detalhada dos arquivos que apresentaram valores m\u00e9dios mais altos e mais baixos dentre os grupos analisados.

As figuras 8 e 9 demonstram que, por vezes, a diferen\u00e7a entre os valores de *pitch* \u00e9 bastante significativa. E essa diferen\u00e7a pode favorecer tanto o diagn\u00f3stico da patologia (caso da figura 8), como tamb\u00e9m pode levar a um resultado err\u00f4neo, como se pode observar na figura 9.

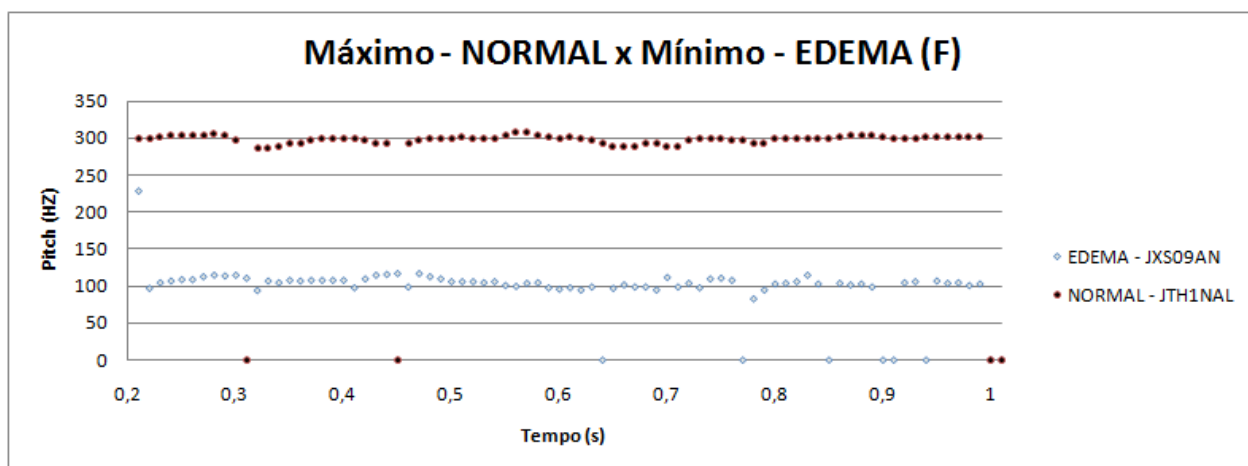


Figura 8. Valores para o pitch em cada segmento de an\u00e1lise em uma amostra de 1s de voz para os arquivos com valores m\u00e9dios m\u00e1ximo e m\u00ednimo - Vozes Femininas (Normal e EDEMA).



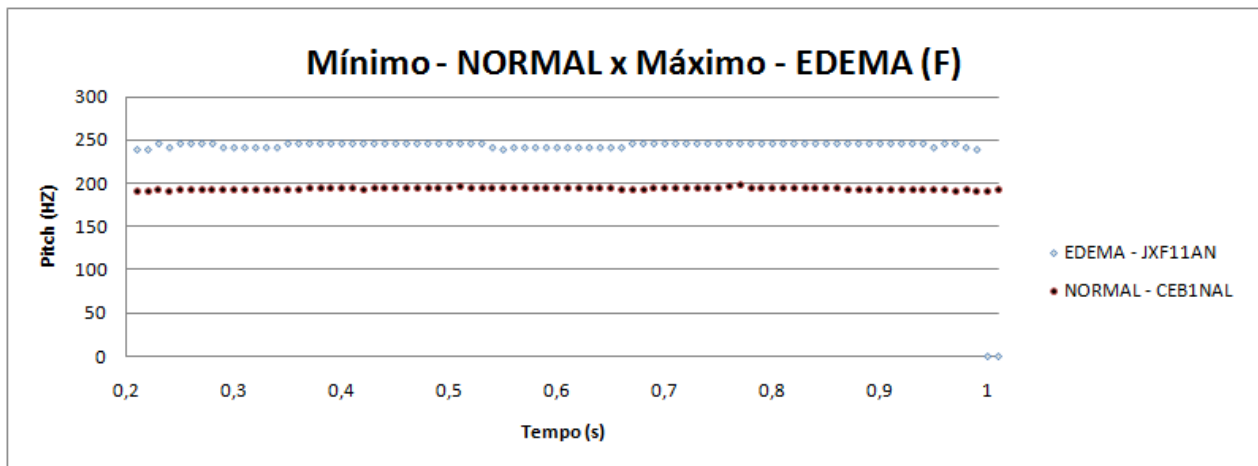


Figura 9. Valores para o pitch em cada segmento de análise em uma amostra de 1s de voz para os arquivos com valores médios mínimo e máximo - Vozes Femininas (Normal e EDEMA).

A análise realizada para o conjunto de parâmetros escolhidos (*jitter*, *shimmer*, quociente de perturbação de amplitude, quociente de perturbação do *pitch*, relação harmônica-ruído e índice de turbulência vocal) resultou nos vários resultados apresentados nos gráficos mostrados a seguir.

Primeiro estão expostos os resultados para o grupo feminino. O *jitter* percentual, o *shimmer* percentual, o quociente de perturbação de amplitude e a relação harmônica-ruído mostraram-se como melhores indicadores da presença da patologia.

As outras características também seguiram o padrão de apresentarem valores mais baixos e com menor variação para vozes normais, porém com uma maior proximidade dos valores obtidos para vozes patológicas quando comparados com aqueles obtidos por meio das características citadas no parágrafo anterior.

Para este primeiro grupo analisado (vozes femininas), os resultados indicam que, quando utilizadas de forma conjunta, as características acústicas podem sim auxiliar no diagnóstico de patologias. Ou seja, as características acústicas destacadas acima podem ser utilizadas para o processo de modelagem paramétrica em um classificador específico, de forma a melhor avaliar e fundamentar a sua importância e o seu comportamento na patologia em estudo.

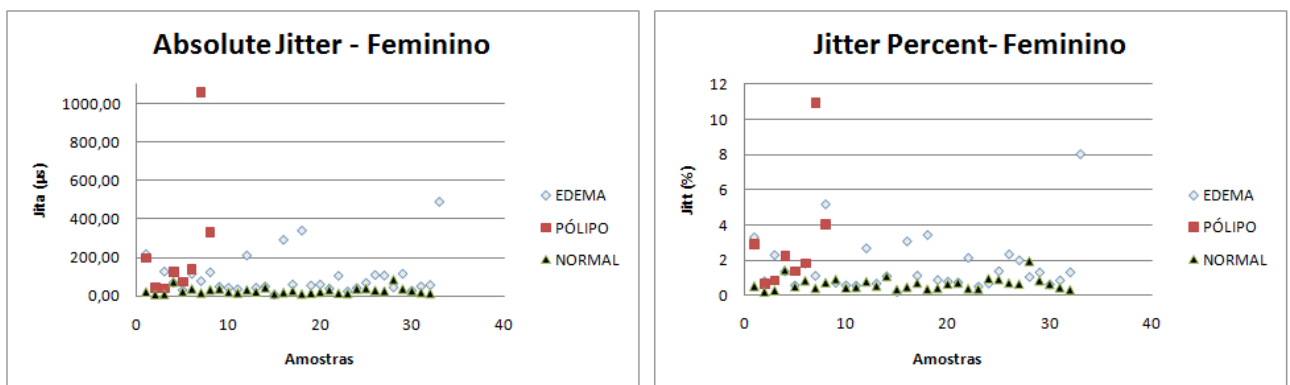


Figura 10. Valores de *jitter* obtidos no software MDVP – Jita (µs) e Jitt(%) - para os arquivos de vozes femininas.

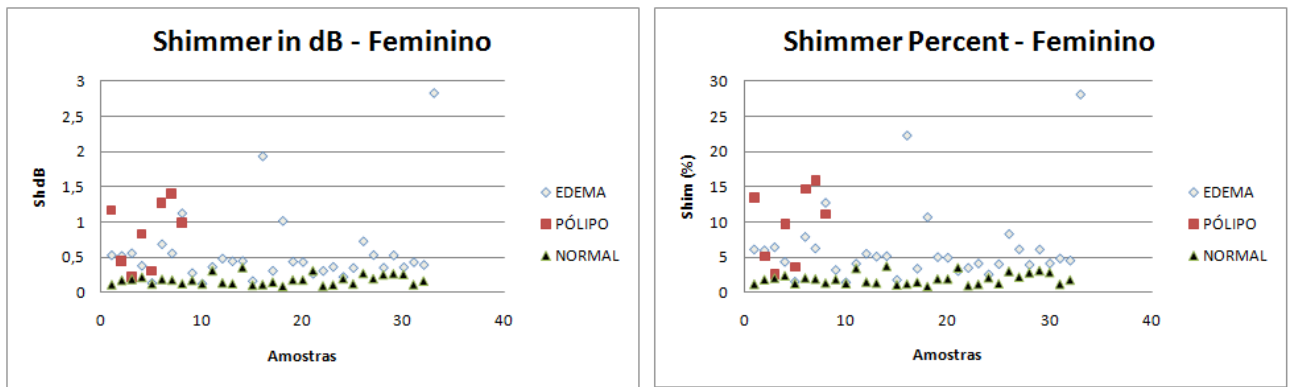


Figura 11. Valores de *shimmer* obtidos no *software* MDVP – ShdB e Shim(%) - para os arquivos de vozes femininas.

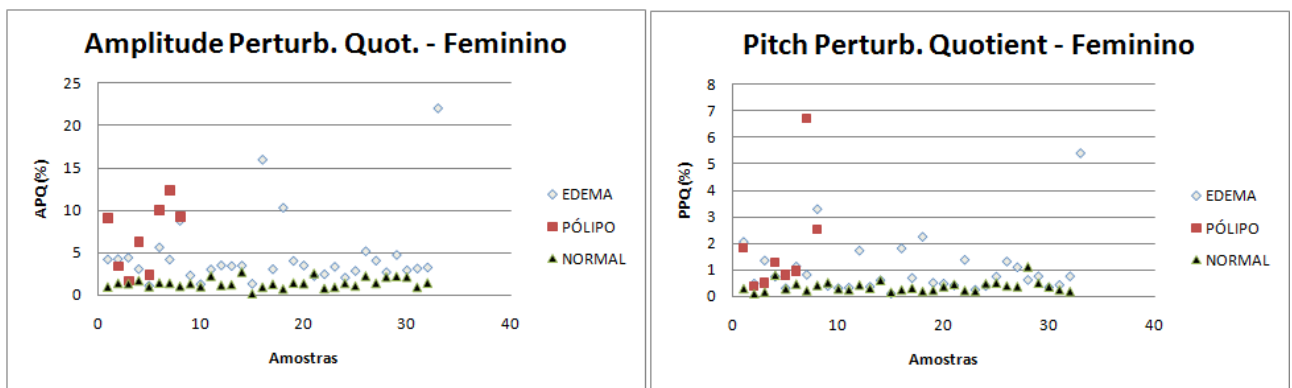


Figura 12. Valores de quociente de perturbação de amplitude e quociente de perturbação do *pitch* obtidos no *software* MDVP – APQ (%) e PPQ(%) - para os arquivos de vozes femininas.

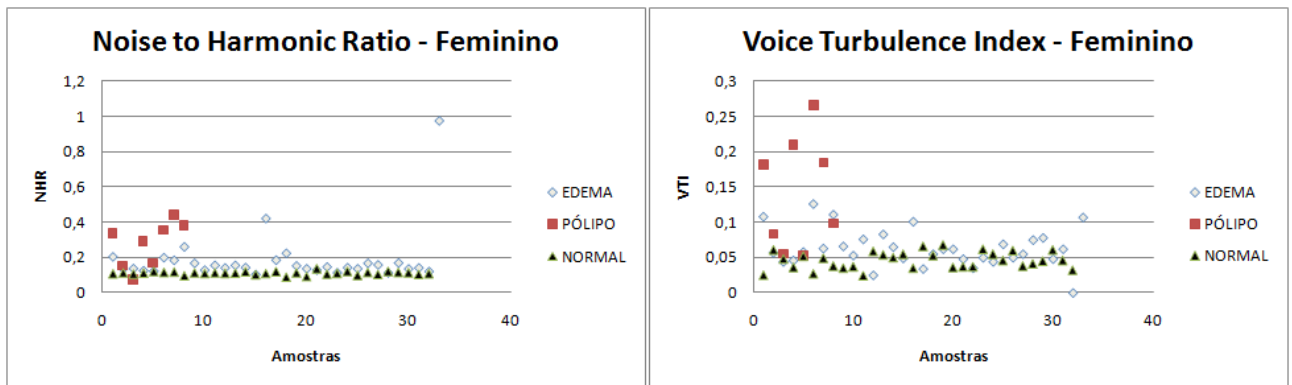


Figura 13. Valores de relação harmônica-ruído e índice de turbulência vocal obtidos no *software* MDVP – NHR e VTI - para os arquivos de vozes femininas

No caso do grupo masculino, os resultados encontrados mostram um menor grau de separação entre vozes patológicas e vozes normais se comparados com os resultados obtidos para o grupo feminino. Tal fato decorre de uma maior faixa de variação de valores para as características dentro das vozes normais.

Nos gráficos abaixo, os triângulos pretos (representam os valores para vozes normais) estão bem mais espaçados verticalmente do que aqueles para o grupo feminino (ver figuras de 10 a 13).

Mesmo com essa diferença, os resultados obtidos para o jitter absoluto, o quociente de perturbação de amplitude e a relação harmônica-ruído mostraram-se bastante satisfatórios individualmente para o processo de discriminação entre vozes normais e as vozes patológicas consideradas.

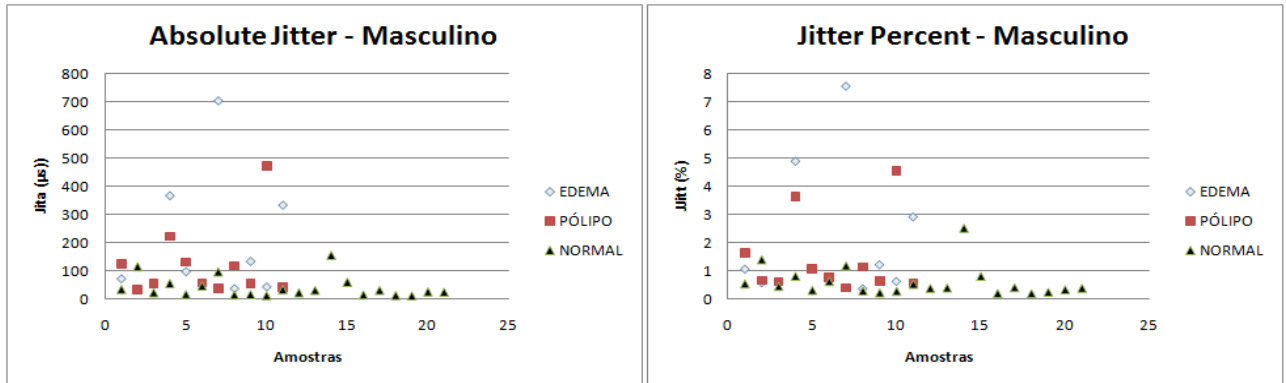


Figura 14. Valores de *jitter* obtidos no *software* MDVP – Jita ( $\mu$ s) e Jitt(%) - para os arquivos de vozes masculinas

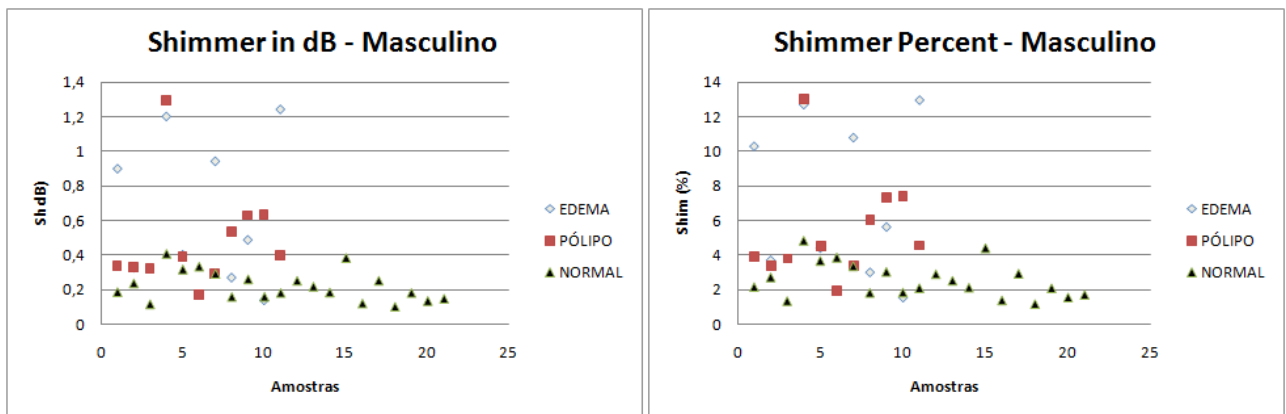


Figura 15. Valores de *shimmer* obtidos no *software* MDVP – ShdB e Shim(%) - para os arquivos de vozes masculinas

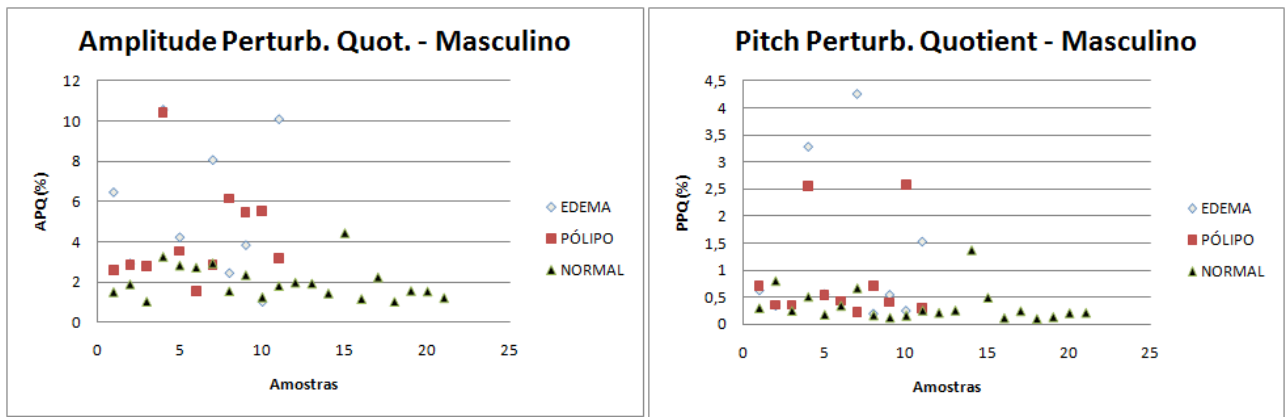


Figura 16. Valores de quociente de perturbação de amplitude e quociente de perturbação do *pitch* obtidos no *software* MDVP – APQ (%) e PPQ(%) - para os arquivos de vozes masculinas

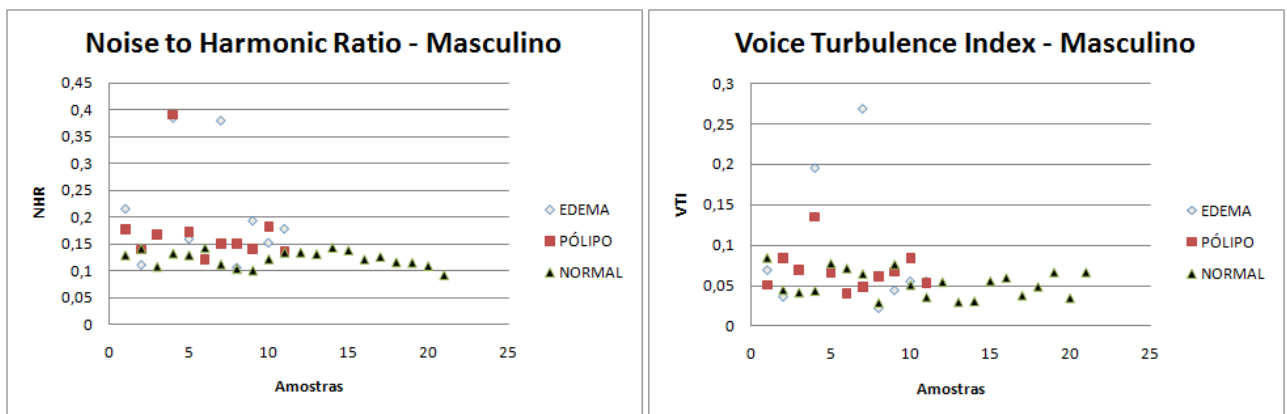


Figura 17. Valores de relação harmônica-ruído e índice de turbulência vocal obtidos no *software* MDVP – NHR e VTI - para os arquivos de vozes femininas

Para evidenciar melhor os resultados obtidos, as tabelas 6 e 7 apresentam os valores médios e de desvios padrão, respectivamente, para cada uma das características escolhidas para análise dentro de cada grupo.

Os valores obtidos para vozes normais, como era esperado, mantiveram o desvio padrão dentro da faixa aceitável. Porém, para as vozes patológicas, observou-se uma variação bastante significativa nos valores para o desvio padrão, o que nos leva a concluir que existem casos de vozes para as quais as características extraídas ficaram bastante distantes dos valores aceitáveis.

Tabela 6: Valores médios para cada característica acústica extraída

|                | Normal<br>(Fem) | Norma<br>(Mas) | Edema<br>(Fem) | Edema<br>(Mas) | Pólipo<br>(Fem) | Pólipo<br>(Mas) |
|----------------|-----------------|----------------|----------------|----------------|-----------------|-----------------|
| Jita( $\mu$ s) | 26,926          | 41,495         | 101,270        | 202,798        | 250,993         | 123,066         |
| Jitt(%)        | 0,635           | 0,589          | 1,681          | 2,238          | 3,113           | 1,414           |
| ShdB           | 0,175           | 0,219          | 0,566          | 0,655          | 0,832           | 0,486           |
| Shim (%)       | 1,994           | 2,524          | 6,308          | 7,261          | 9,516           | 5,388           |
| APQ (%)        | 1,361           | 1,995          | 4,568          | 5,519          | 6,795           | 4,268           |
| PPQ (%)        | 0,367           | 0,339          | 1,043          | 1,286          | 1,876           | 0,834           |
| NHR            | 0,112           | 0,123          | 0,184          | 0,209          | 0,276           | 0,176           |
| VTI            | 0,046           | 0,053          | 0,107          | 0,092          | 0,141           | 0,069           |

Tabela 7: Valores para o Desvio Padrão em cada característica extraída

|                | Normal<br>(Fem) | Norma<br>(Mas) | Edema<br>(Fem) | Edema<br>(Mas) | Pólipo<br>(Fem) | Pólipo<br>(Mas) |
|----------------|-----------------|----------------|----------------|----------------|-----------------|-----------------|
| Jita( $\mu$ s) | 16,922          | 37,399         | 102,327        | 225,457        | 340,250         | 129,451         |
| Jitt(%)        | 0,356           | 0,547          | 1,569          | 2,467          | 3,350           | 1,375           |
| ShdB           | 0,071           | 0,087          | 0,526          | 0,419          | 0,455           | 0,305           |
| Shim (%)       | 0,805           | 1,022          | 5,441          | 4,447          | 5,156           | 3,051           |
| APQ (%)        | 0,581           | 0,852          | 4,231          | 3,441          | 3,950           | 2,496           |
| PPQ (%)        | 0,208           | 0,297          | 1,050          | 1,482          | 2,082           | 0,871           |
| NHR            | 0,009           | 0,015          | 0,154          | 0,104          | 0,129           | 0,074           |
| VTI            | 0,012           | 0,017          | 0,024          | 0,083          | 0,079           | 0,026           |

Como observação final, vale salientar que o conhecimento prévio do sexo do paciente a ser avaliado, bem como hábitos de vida, ser fumante ou não-fumante por exemplo, são de certa forma imprescindíveis para uma análise mais acertada. Pode-se observar que uma voz patológica feminina apresenta algumas características com valores próximos daqueles apresentados por vozes masculinas saudáveis.

De maneira geral, os resultados demonstram que as características acústicas destacadas poderão ser utilizadas para o processo de modelagem acústica para utilização em classificadores específicos, de forma a melhor avaliar e fundamentar a sua importância e o seu comportamento na patologia em estudo.

Tal tarefa constituiria a próxima etapa de pesquisa para a continuação e utilização dos dados obtidos.

## CONCLUSÕES

Os resultados mostraram que pode ser comprovada a alteração de características acústicas da voz quando da presença de patologias. Ficou evidente por meio da pesquisa realizada que a modificação dos pulsos glotais e a passagem destes pelo aparelho fonador refletem tais alterações.

A frequência fundamental, parâmetro importante para a caracterização de qualquer sinal, inclusive o de voz, é um indicador que deve ser manipulado com atenção. Sua utilização individual não leva a resultados precisos. Porém quando a tarefa de sua extração é feita de maneira acertada e com a maior precisão possível, sua utilização em conjunto com outras características clássicas, algumas das quais apresentando significativa dependência da própria frequência fundamental, esta mostra-se como alternativa promissora no diagnóstico de patologias na laringe.

A tarefa de extração dos parâmetros em vozes muito ruidosas mostrou-se uma tarefa extremamente complicada, muitas vezes impossível de ser realizada ou então retornando valores totalmente fora da faixa. O que pode levar a erros em análises estatísticas.

## AGRADECIMENTOS

Ao CNPq pelo financiamento do projeto e pela concessão da bolsa PIBIC;

A todos os colegas do Laboratório de Automação e Processamento de Sinais (LAPS) da UFCG pela disponibilidade em ajudar com a configuração das máquinas e a toda equipe que contribuiu com o desenvolvimento deste projeto.

## REFERÊNCIAS BIBLIOGRÁFICAS

S. B. Davis, "Acoustic Characteristics of Normal and Pathological Voices", *Speech and Language: Advances in Basic Research and Practice*, Vol. 1, pp. 271–335, 1979.

J. I. Godino-Llorente, P. Gomes-Vilda and M. Blanco-Velasco, "Dimensionality Reduction of a Pathological Voice Quality Assessment System Based on Gaussian Mixture Models and Short-Term Cepstral Parameters", *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 10, pp. 1943-1953, October, 2006.

V. Parsa and D.G. Jamieson, "Acoustic Discrimination of Pathological Voice: Sustained Vowels Versus Continuous Speech", *Journal of Speech, Language, and Hearing Research*, Vol. 44, pp. 327-339, April 2001.

L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signal*. New Jersey: Prentice Hall, 1978. Kay Elemetrics Corp. Disordered Voice Database, Model 4337, 03 Ed, 1994.

B. G. Aguiar Neto, Silvana C. Costa, Joseana M. Fachine, Menaka Muppa. "Feature Estimation for Vocal Fold Edema Detection Using Short-Term Cepstral Analysis". *IEEE 7<sup>th</sup> International Symposium on BioInformatics & BioEngineering*, Boston, USA, October 2007

J. I. Godino-Llorente and P. Gómez-Vilda. "Automatic Detection of Voice Impairments by Means of Short-Term Cepstral Parameters and Neural Network Based Detectors". *IEEE Trans. on Biomedical Engineering*, Vol. 51. No. 2, pp. 380-384, February 2004.

David Gerhard, "Pitch Extraction and Fundamental Frequency: History and Current Techniques", Technical Report, TR-CS 2003-06, November, 2003.

J. I. Gondino-Llorent, "On the selection of meaningful speech parameters used by pathologic/non pathologic voice register classifier". *Proceedings of Eurospeech'99*, Budapest, Hungary, 1999, pp 563-566.

E. Perrin, C. Berger-Vachon, I. Kauffmann, and L. Collet, "Acoustical recognition of laryngeal Pathology using the fundamental frequency and the first three formants of vowels", vol. 35, no. 4, pp.361–368, 2006.

B. Boyanov and S. Hadjitodorov. "Acoustic Analysis of Pathological Voices". *IEEE Engineering in Medicine and Biology*, pp. 74-82, July/Agust, 1997.

J. Min Choi, Myung-Whun Sung and K. Suk Park, "New Method in Acoustic Analysis for the Diagnosis of the Laryngeal Functions", *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, TX, USA, October 23-26, 2002.

Silvana C. Costa, "Análise Acústica, Baseada no Modelo Linear de Produção da Fala, para Discriminação de Vozes Patológicas", Tese de Doutorado, Universidade Federal de Campina Grande, Novembro, 2008.

C. E Martinez and H. L. RUFINER, "Acoustic Analysis of Speech for Detection of laryngeal Pathologies", *Proceedings of the 22th Annual EMBS Conference*, pages 2369-2372, Chicago, July, 2000.

Taherkhani, A., Seyyedsalehi, S.A., Mohammadi, A., Moradi, M.H.,” Nonlinear Signal Processing for Voice Disorder Detection by Using Modified GP Algorithm and Surrogate Data Analysis”, 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 15-18 December, . 2007.

P. Gómez-Vilda, R. Fernández-Baillo, V. Rodellar-Biarge, V. N. Lluís, A. Álvarez-Marquina, L. M. Mazaira-Fernández, R. Martínez-Olalla, J. I. Godino-Llorente, “Glottal Source biometrical signature for voice pathology detection”, *Speech Communication*, Volume 51, Pages 759-781, September 2009.